猪三维点云体尺自动计算模型Pig Back Transformer

王宇啸[2]，石源源[1]，陈招达[1]，吴珍芳[2,3,4]，蔡更元[2,3,4]，张素敏[1]，尹　令[1,2,4*]

(1. 华南农业大学 数学与信息学院,广东广州 510642,中国；2. 国家生猪种业工程技术研究中心,广东广州 510642,中国；3. 华南农业大学 动物科学学院,广东广州 510642,中国；4. 猪禽种业全国重点实验室,广东广州 510640,中国)

摘　要：［**目的/意义**］为了提高体尺关键点定位准确率，猪三维点云体尺自动测量方法会采用点云分割，在各个分割后局部点云定位测量关键点，以减少点云之间相互干扰。然而点云分割网络通常需要消耗较大计算资源，且现有测量点定位效果仍有待提升空间。本研究旨在通过设计关键点生成网络从猪体点云中提取出各体尺测量所需关键点。在降低显存资源需求的同时提高测量关键点定位效果，提高体尺测量的效率和精度。［**方法**］针对猪三维表面点云进行体尺测量，提出了一种定位猪体尺关键点的模型Pig Back Transformer。模型分为两个模块，分别设计了两种改进的Transformer自注意力编码器，第一模块为全局关键点模块，首先设计了一种猪背部边缘点提取算法用于获取边缘点，再使用edge encoder编码器以边缘点集合作为输入，edge encoder的edge attention中加入了边缘点和质点的偏移距离信息；第二模块为关键点生成模块，使用了back attention机制的back encoder，其中加入了与质心和第一模块生成的全局关键点的偏移量，并将偏移量与点云注意力通过按位max pooling操作结合，最后通过生成猪的体尺测量关键点和背脊走向点。最后设计了使用关键点和背脊走向点作为输入的体尺算法。［**结果和讨论**］对比关键点和背脊走向点生成任务上Pig Back Transformer表现最佳，并对比体尺计算结果与人工测量结果，体长相对误差为0.63%，相对PointNet++、Point Transformer V2、Point Cloud Transforme、OctFormer PointTr等模型有较大提升。［**结论**］Pig Back Transformer能相对准确地生成猪体尺关键点，提高体尺测量数据准确度，并且通过点云特征定位体尺关键点节省了计算资源，为无接触牲畜体尺测量提供了新思路。

关键词：Pig Back Transformer；三维点云；体尺自动测量；测量关键点定位；深度相机；自注意力机制

中图分类号：S2　　**文献标志码：**A　　**文章编号：**SA202401023

引用格式：王宇啸, 石源源, 陈招达, 吴珍芳, 蔡更元, 张素敏, 尹令. 猪三维点云体尺自动计算模型Pig Back Transformer[J]. 智慧农业(中英文), 2024, 6(4): 76-90. DOI：10.12133/j.smartag.SA202401023

WANG Yuxiao, SHI Yuanyuan, CHEN Zhaoda, WU Zhenfang, CAI Gengyuan, ZHANG Sumin, YIN Ling. Pig Back Transformer: Automatic 3D Pig Body Measurement Model[J]. Smart Agriculture, 2024, 6(4): 76-90. DOI：10.12133/j.smartag.SA202401023 (in Chinese with English abstract)

0　引　言

培育优良种猪和控制生猪最佳出栏时间是畜牧业关注的重要问题，特定时间段的猪表型体尺测量是种猪培育和生猪出栏的关键参考点。表型体尺量化地描述了猪体外部特征、表型特征、发育情况等。体尺表型值常可作为育种时的遗传参数，亦可以作为育肥潜力的参考用以控制最佳出栏时间。目前猪表型体尺数据获取的常规方法主要是人工测量法，然而人工测量费时费力、效率低下、同时结果受到测量人员专业素养的主观性影响，且测量过程容易引发牲畜应激反应，并有可能对测量人员造成人身伤害，目前大规模人工测量已经不能满足现代大型养殖场的需求。

随着计算机视觉的发展，牲畜体尺的无接触测量方法越来越引起重视，其技术也经历过了多次改变和发展。最初的研究是使用彩色相机获取猪单视角下的二维彩色图像，通过对二维图像的分析和计算获取牲畜体尺或进行体重估计。江杰等[1]运用基于灰度的背景差分法，结合色度不变性原理，从

收稿日期：2024-01-21

基金项目：国家自然科学基金面上基金（32172780）；国家重点研发计划子课题 （2023YFD1300202）

作者简介：王宇啸，研究方向为计算机视觉。E-mail： wangyuxiaojoemob@stu.scau.edu.cn

*通信作者：尹　令，博士，教授，研究方向为三维视觉。E-mail： yin_ling@scau.edu.cn

复杂环境中检测羊体提取轮廓包络线，并获取轮廓包络线最大曲率点作为羊体臀部测点，采用四点法求取羊体其他3个测点计算出羊体的体尺参数。Zhang等[2]在获取羊只顶部和侧面图像后，使用SLIC（Simple Linear Iterative Clustering）Superpixel算法通过分割提取图像中的羊体，再分析轮廓包络线完成羊的体长、体高等体尺信息估算。

使用二维视觉进行采集的无接触测量方法后续探索中开始出现采用二维图像三维重建方法的研究。Yang等[3]通过使用二维图像进行三维重建得到点云，再进行奶牛的体尺测量；奶牛在测量过程中颈部被固定在围栏上，从奶牛右侧沿背后至左侧连续采集图像，使用SIFT（Scale-Invariant Feature Transform）算法[4]进行图像关键点提取，通过关键点确定相机位置参数，使用位置参数重建点云，经过点云提取、姿态归一化和点云补全后，通过聚类进行定位并计算体尺。

二维图像估算牲畜体尺具有成本低、设备安装简易等优势，然而二维图像缺少深度维度信息，且相机在物距、光照条件改变时，测量结果都会受到影响，故通用性较差，难以提升系统整体准确性。

随着消费级深度相机的广泛应用，研究者逐渐转向使用深度相机获取牲畜的三维表面点云进行体尺测量和体重估计。深度相机获取的点云数据包含了X、Y、Z坐标和颜色等信息，拓展了牲畜体尺测量的深度和广度研究。Salau等[5]使用单视角深度相机获取奶牛背部三维点云，分析统计采集点云定位体尺关键点实现对奶牛体尺自动测量，受到单视角的局限，该方法只能对背部和臀部部分体尺进行测量。叶文帅等[6]通过分析肉牛骨架特征和肉牛图像边缘轮廓特征,提出一种多姿态肉牛体尺自动测量方法。Ruchay等[7]使用非刚性配准提高配准精度，并通过在牛体进行物理标记的方法在采集的图像上得到关键点进行测量。Shi等[8]提出了一种基于注意力的特征提取模型对奶牛进行了体况评分，并在体况评分任务中取得了较好的结果；同时构建了一个由512头奶牛构成的数据集。Hu等[9]尝试使用改进的PointNet++网络对点云做分割，平均准确度为标准姿态下96.41%，非标准姿态下94.98%，对分割后的点云进行关键点定位，点云各部分分割后对猪的姿态变化有更高的容忍度，泛化性有所提升，但受限于定位点的精度和定位点数量，仍然有提升空间。

综上所述，采用牲畜表面点云进行体尺计算中，牲畜点云姿态多种多样，会呈现出不同的局部特征，使得体尺测量中关键点定位算法难以泛化应用于全部点云中[10]，这也使得如何能更精确地通过点云对牲畜进行体尺测量成为了一个难点。

采用消费级深度相机如Azure Kinect等采集的点云具有非结构化、无序的特点，目前点云处理任务如分割、配准、补全、目标检测等多采用神经网络实现，其发展从早期基于体素的卷积神经网络，之后的基于点的卷积神经网络，如PointNet[11]，到目前兴起的应用于点云的使用自注意力机制的神经网络如Point Transformer[12]，Fast Point Transformer[13]，Point-BERT[14]。目前的网络设计目的多为分割和分类，这些网络模型有着优秀的提取点云特征的能力，可以修改后作为任务的主干网络。

为进一步提升猪体尺测量的精度，且不采用消耗大量计算资源的点云分割模型，本研究设计了一种猪体点云自注意力机制网络模型Pig Back Transformer进行测量关键点定位，并使用关键点进行体长、体高、体宽、腹围等体尺数据的自动测算。

1 研究方法

本研究猪点云背部关键点定位及体尺测量算法计算流程如图1所示。

采集和预处理模块：首先使用深度相机采集顶部、左侧、右侧三个视角的局部点云；再分别将各视角点云中的背景去除，并对局部点云配准后得到完整三维点云；之后进行去噪、平滑处理和下采样；最后使用背部点云边缘提取算法获得背部边缘点云。

体尺计算模块：背部边缘点和内部点通过Pig Back Transformer模型生成背部测量关键点，使用背部测量关键点和背脊走向点结合背部点云计算体长，结合整体点云计算体高、体宽、腹围、胸围。

1.1 采集和数据预处理

本研究的数据采集地点位于广东省温氏河源新晶公司，采集猪种包括长白猪、大白猪、杜洛克、皮特兰和蓝塘土猪，猪龄为160～220天。分两次进行数据采集。第一次于2020年7月开始，持续30天；第二次于2022年7月开始，持续45天。

第一次采集，猪只体重为67.79～142.67 kg，平均体重105.28 kg，体长为100～130 cm，平均体长115.79 cm，体高为56～70 cm；第二次采集时，猪只体重为60～150 kg，体长为80～140 cm，体高为50～90 cm。

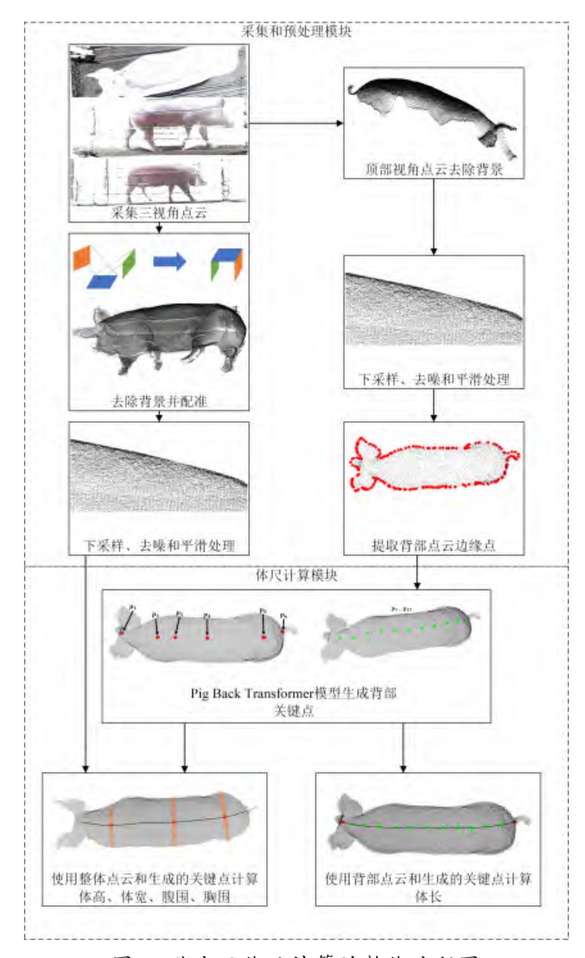


图 1 猪点云体尺计算的整体流程图

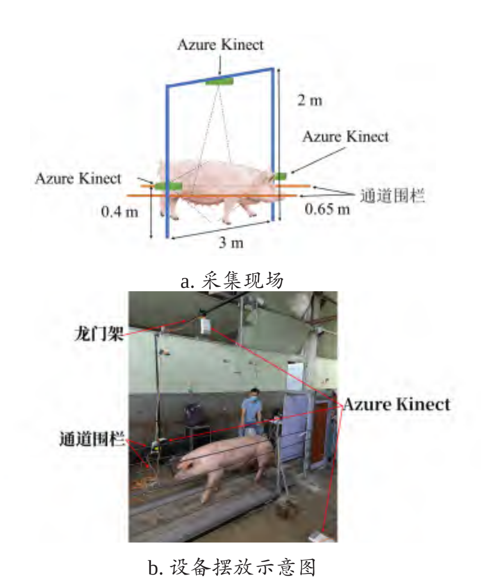
Fig. 1 Process flow diagram of pig's point cloud body size calculation

采集通道由两侧各三根钢制栏杆构成，宽度为0.65 m，长度为2.5 m。

第一次采集时使用三台Kinect V2相机从通道顶部以及两侧进行拍摄，相机参数为：深度图像分辨率像素为512×424，水平视角70°，垂直视角60°，深度范围0.5～3.5 m。通道顶部相机位于中央正上方高度为2.2 m处，两侧相机与地面相对高度为40 cm，水平间距为4.5 m。

第二次采集时使用三台Azure Kinect相机从通道顶部以及两侧进行拍摄，相机参数为深度图像分辨率像素为1 024×1 024，水平视角120°，垂直视角120°，深度范围0.25～2.21 m。通道顶部相机位于中央正上方高度为2 m处，两侧相机与地面相对高度为40 cm，水平间距为3 m，此次采集现场如图 2a所示，采集设备摆放示意图如图 2b所示。

第一次采集得到30头猪的264组点云，去除掉不完整点云以及样本量过少的点云后，得到20头猪，每头猪10组点云，共200组点云，每组都包括了背部、两侧和整体点云，并手工测量了所有受采集猪只的体尺长度数据；第二次采集得到120头猪的1 022组点云，在去除掉不完整点云后，保留了70头猪的610组点云，每组都包括了背部、两侧和整体点云。



a. 采集现场

b. 设备摆放示意图

图 2 使用Azure Kinect的猪体点云采集通道现场与设备摆放示意图

Fig. 2 Tunnel for pig's point cloud collect using Azure Kinect

研究中进行了边缘提取、神经网络性能和体尺计算三部分实验，边缘提取实验数据集共810个猪体背部点云；关键点生成实验中，人工标记了810组点云作为真值，每组人工标记包括体长测量起始点、前体高和肩宽测量关键点、胸围测量关键点、腹围和腹宽测量关键点、后体高和臀宽测量关键点、体长测量终点以及24～58个根据体长变化的背脊走向点，实验输入的数据集为810个背部点云、各自的边缘点和人工标记的真值经过对称、旋转操作进行数据增强得到的共4 050组数据，每组包括一个背部点云和此背部点云的边缘点，训练时采用随机分割策略，数据分为8∶1∶1的比例分别用作训练集、验证集、测试集；体尺计算实验数据集选取了人工测量过的20头猪，每头10个背部点云、10个完整点云，共400个点云；体长分布实验数据集为从体尺计算实验数据集中根据背脊走向曲线的曲率取姿态变化剧烈的10头，每头10个背部点云，10个完整点云，共200个点云。

1.1.1 配准并提取猪体点云

采集到的点云数据包含有背景，需要提取出猪体目标点云。提取步骤为分别采集通道背景点云集

合S_b，及含有猪体的点云S_a，猪体目标点云为S_p，阈值为k_p，如公式（1）所示，根据距离背景点云的距离确定是否为猪体，与背景距离大于阈值的点为猪体点云，时间复杂度为$o(mn)$，m为猪体点数，n为总点数。

$$S_p = \left\{ \forall p \in S_a | \nexists q \in S_b, \left|\left| p, q \right|\right|_2 < k_p \right\} \tag{1}$$

采集所用的3台深度相机从不同视角获取的猪体局部点云位于各自坐标系中，需要统一到同一视角坐标系下。本研究采用的配准方法如图3所示，设顶部相机为世界坐标系，采用刚性立方体配准物将左侧和右侧摄像机的点云变换到世界坐标系下，详细内容参见文献[15]。



图3　Pig Back Transformer模型的刚性立方体配准物配准过程

Fig. 3 Registration process of the rigid registration box in Pig Back Transformer model

1.1.2　噪声点去除

猪体目标点云附着有大量噪声，猪体点云近点噪声由猪体与栏杆之间光线的反复衍射和环境中灰尘等因素产生，噪声点会对边缘提取和关键点生成产生极大的影响。本研究采用统计滤波和半径滤波组合的两级滤波器对猪体点云进行去噪。

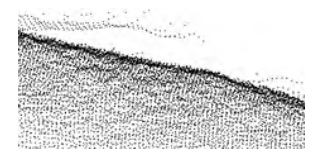
统计滤波用于去除离群点，其原理为，设置邻域点数量和标准差，设每个点与邻近点之间的距离服从高斯分布，平均距离为均值，可以根据均值和每一个距离求得标准差，平均距离在标准范围之外的点可以被定义为离群点，并从数据中去除。

半径滤波器则是用于去除小于猪体点云密度的噪声点云。半径滤波器以点云中某一点为中心画一个圆计算落在该圆中点的数量，当数量大于邻域点阈值k，则保留该点，否则剔除该点。

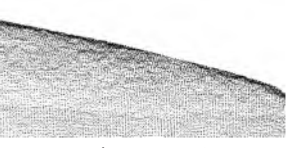
1.1.3　猪体点云平滑

点云经过去噪后，远点噪声和大部分近点噪声基本被去除，但是贴近猪体表面的部分毛刺噪声点（图4a）依旧没有去除。这些毛刺噪声会造成猪体点云表面出现锐利区域，使后续的点云法向量估计出现偏差，造成边缘点提取错误。

为了解决上述问题，采用局部范围内移动最小二乘法[16]将猪体表面点云拟合为平滑的曲面。点云平滑效果如图4b所示，基本消除了表面的毛刺噪声。



a. 表面毛刺点



b. 平滑后的猪体点云

图4　使用局部范围内移动最小二乘法的点云平滑效果

Fig. 4 The smoothing effect of point cloud using local range moving least squares method

1.1.4　猪背部边缘提取

提取猪背部轮廓边缘的目的是获得更加稳定的全局特征，在使用模型提取点云全局几何特征时，三维相机的成像原理导致大量点云集中在背部的平坦区域。平坦区域点在神经网络中提供的特征是相似的，全部输入网络将造成计算资源浪费，故本研究选择采样背部边缘点作为模型第一部分的输入，不仅可以获得稳定的全局特征，而且能极大地减少模型所需的参数量。

背部点云边缘提取算法先根据质点修整点云，再计算每个点的k点法向量并拟合法平面，最后对每个点的k点在其法平面上的投影与其构成的邻边夹角判断是否为边缘点。

顶部相机获取猪体点云时，由于视角问题会导致如图5a所示的边缘不平整现象，直接进行边缘采样会影响模型的稳定性及拟合速度降低。本研究将背部点云进行切片，并计算切片质心，将质心以下的点舍去。修整后的点云如图5b所示，不仅仍然能提供完整的全局特征，同时还可以使用更少的点得到稳定的全局特征。



a. 未修剪的猪体背部点云　　b. 修剪后的点云

图5　修剪前后的猪体背部点云对比

Fig. 5 Origin point cloud & trimmed point cloud

计算法平面时，先使用k-d树构建索引并确定每个点的k近邻点，并根据k近邻点使用最小二乘法估计法平面。

在计算得到点云中每一点的法平面后，还需计算点云中每一点和其最近点在法平面上的投影构成的邻边夹角。由于点云具有无序性，在直接计算邻边夹角时需要先进行投影，再经过$o(n^2)$级时间复杂度的计算才能确定每个邻边夹角，会额外消耗大

量计算资源。

通过三个变换矩阵使当前进行判断的点位于原点，并使法平面与xOy平面重合，得到变换后的最近点$\boldsymbol{Z}_k = \boldsymbol{T}_3\boldsymbol{T}_2\boldsymbol{T}_1\boldsymbol{X}_k$，将$\boldsymbol{Z}_k$投影到$xOy$平面后计算极角，再利用极角计算邻边夹角以降低算法的复杂度，变换过程如图6所示。

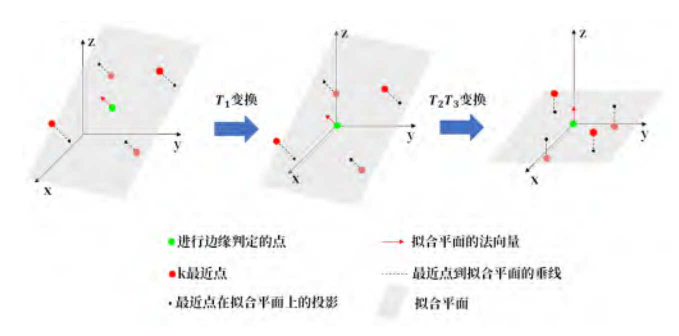


图6 边缘点提取中的拟合平面的变换过程

Fig. 6 The transformation process of the fitted plane in edge point extraction

图6中的$\boldsymbol{T}_1$，$\boldsymbol{T}_2$，$\boldsymbol{T}_3$如公式（2）～公式（4）所示。

$$\boldsymbol{T}_1 = \begin{bmatrix} 1 & 0 & 0 & -\bar{p}_1 \\ 0 & 1 & 0 & -\bar{p}_2 \\ 0 & 0 & 1 & -\bar{p}_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

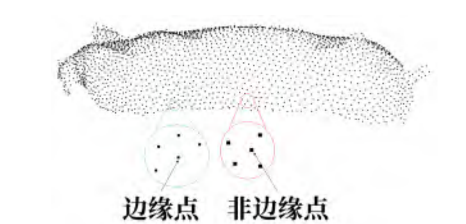
$$\boldsymbol{T}_2 = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$$\boldsymbol{T}_3 = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

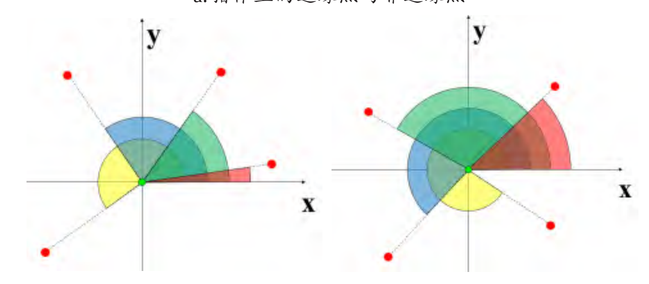
式中：$\boldsymbol{T}_1$为将A平移到原点的刚性变换；$\boldsymbol{T}_2$为绕Z轴的旋转变换；α为法向量在xOy平面的投影和x轴的夹角；$\boldsymbol{T}_3$为绕Y轴的旋转变换；β为进行$\boldsymbol{T}_2$变换后法向量与z轴的夹角；由于法平面与xOy重合，此时投影可以直接使用$\boldsymbol{Z}_k$的x，y坐标，此方法相比直接投影到法平面再计算极角拥有更低的时间复杂度。

将$\boldsymbol{Z}_k$投影到xOy平面并转换为极坐标，得到一组极角ρ_k，将极角按从小到大排序，相邻极角差即为投影后的邻边夹角，投影效果如图7所示，红色、绿色、蓝色、黄色依次为被判别点的各近邻点投影到最小二乘法平面后的极角。

得到极角并从大到小排序后，依次计算极角差得到邻边夹角。



a. 猪体上的边缘点与非边缘点



b. 边缘点和邻近点的投影　c. 非边缘点和邻近点的投影

图7 边缘提取算法中k近邻在临近点拟合平面上投影的极角

Fig. 7 The projectiles of k nearest in polar coordinate system in edge detection algorithm

设邻边夹角为R_k，则邻边夹角计算如公式（5）和公式（6）所示。

$$R_i = \rho_{i+1} - \rho_i \tag{5}$$

$$R_k = \rho_1 - \rho_k + 2\pi \tag{6}$$

式中：R为邻边夹角，ρ为极角，且$(i = 1, 2, 3, \cdots, k-1)$。此时设阈值为$\gamma$，$\max(R_k) > \gamma$时，当前中心点为边缘点。

1.2 体尺计算

本研究设计的体尺算依托于Pig Back Transformer模型获得的体尺关键点和背脊走向点，体尺关键点协助定位各体尺计算区域，背脊走向点曲线拟合可以进行体长计算。

1.2.1 改进的自注意力Pig Back Transformer

猪体背部点云轮廓边缘提取模型是在Transformer网络基础上提出了图8所示的Pig Back Transformer网络结构。首先进行点云归一化，之后将边缘点云使用最远点采样到32个点作为输入，经Edge Embedding映射，并通过Edge Transformer模块提取特征后由多层感知机得到全局特征点，再输入经过尺寸为0.01的体素降采样和512个点的最远点采样后的完整背部点云，并与全局特征点一起通过Back Embedding完成嵌入，再使用Back Transformer模块提取特征，最后将特征通过多层感知机得到相对偏移，使用偏移和全局特征点求和得到关键点。

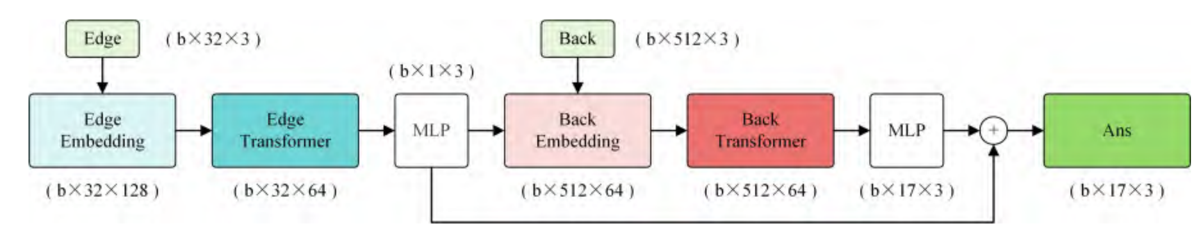


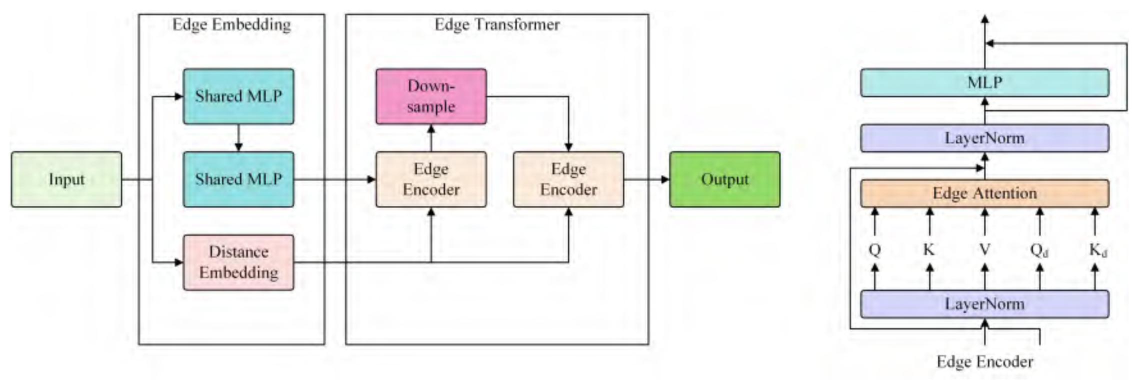
图 8　本研究提出的Pig Back Transformer结构

Fig. 8　The Pig Back Transformer's structure proposed in this research

下文依次介绍Pig Back Transformer各组件的设计。

1）改进嵌入Edge Embedding和自注意力Edge Transformer。

Edge Embedding结构如图9所示，分为两部分。第一部分使用PointNet++ [17] 中的多层感知机Shared MLP（Multilayer Perceptron）进行映射，Shared MLP使用相同的线性变换将点云中的所有点映射到相同的高维空间中，先经过输入为3个维度，输出为128个维度的卷积，再经过输入为128个维度，输出为128个维度的卷积得到高维点特征$\boldsymbol{X}_p$，$\boldsymbol{X}_p$将作为edge attention中生成Query（$\boldsymbol{Q}$）、key（$\boldsymbol{K}$）、Value（$\boldsymbol{V}$）矩阵所用的输入；第二部分使用每个点与质心的距离构成向量组作为距离嵌入X_d，X_d为edge attention中生成$\boldsymbol{Q}_d$，$\boldsymbol{K}_d$所用的输入。在Edge Embedding中添加距离嵌入$\boldsymbol{X}_d$的原因是相对质心的距离蕴含了点云的几何信息，在获取全局关键点时，自注意力可以根据$\boldsymbol{X}_d$为不同的点分配权值，结合使用点云本身的三维位置信息得到的注意力权重，可以使最终得到的注意力权重同时利用到X_d的全局几何信息和从三维位置信息提取出的局部特征。



a. Edge Embedding & Edge Transformer　　b. Edge Encoder

图 9　Edge Embedding & Edge Transformer结构

Fig. 9　Structures of Edge Embedding & Edge Transformer

Edge Embedding与Transformer采用了不同的设计。首先没有采用与Transformer中相同的位置编码形式，因为Transformer在Embedding步骤中直接将位置信息添加到了输入的向量中是为了解决self-attention中基于特征提取时会忽略序列顺序导致的词序问题，但是self-attention特征提取的序列无关性却是适合解决点云无序性的关键，所以在进行网络改进的过程中，本文研究选择去除掉位置编码，直接进行特征提取。

其次，Edge Embedding中使用Shared MLP进行提取的设计相对于使用MLP减少了大量的训练参数，Shared MLP为输入点云集合$\boldsymbol{P}$训练矩阵$\boldsymbol{S}$进行映射，得到特征$\boldsymbol{F} = \boldsymbol{PS}$，此时$P$的参数量为$3 \times f$，$f$为特征的维度，若使用MLP进行特征提取，则参数量为$3 \times f \times n$。MLP相对Shared MLP的优点是在进行信息提取时可以将输入进行交叉，而Shared MLP的信息提取是孤立的，只能提取到每个输入自身的高维信息，但是，由于在网络的后续设计中采用了Edge Transformer通过edge attention提取局部和全局信息，所以选择使用Shared MLP作为Embedding步骤的特征提取器。

Edge Transformer使用两层编码器进行一次降采样的结构，编码器中使用了改进的edge attention，其计算如公式（7）所示。

$$Attention\left(X_p, X_d\right) = sm\left(sm\left(\frac{\left(X_p W_{pq}\right)\left(X_p W_{pk}^{\mathrm{T}}\right)}{\sqrt{D_p}}\right) + sm\left(\frac{\left(X_d W_{dq}\right)\left(X_d W_{dk}^{\mathrm{T}}\right)}{\sqrt{D_d}}\right)\right) \times \left(X_p W_v\right) \tag{7}$$

式中：X_p，X_d分别为使用Shared MLP和距离嵌入得到的特征；W为权值矩阵；D为维度。

改进的自注意力机制添加了与质点之间的距离信息，补充了点的相对位置特征信息。通过叠加点特征和距离特征产生的自注意力权值，能得到更具有代表性的全局特征。

Edge Transformer后使用两层的MLP将得到的特征转化为全局特征点。

2）改进嵌入Back Embedding和自注意力Back Transformer。

Back Embedding如图10所示，分为三部分，第一部分是以背部点与全局特征点作差，得到的向量组作为基于全集特征点的偏移向量组，第二部分是将偏移向量组经过线性映射至64维得到X_p，第三部分是使用背部点和偏移向量组构建距离嵌入得到X_d，距离嵌入的算法为：

设3行n列的背部点云向量组为M，3行n列的偏移向量组为$\hat{M}$，质点为G。

X_d的第一维度表示为公式（8）。

$$X_{d1,i} = \hat{m}_{i2} = \sqrt{\hat{M}_{1,i}^2 + \hat{M}_{2,i}^2 + \hat{M}_{2,i}^2} \tag{8}$$

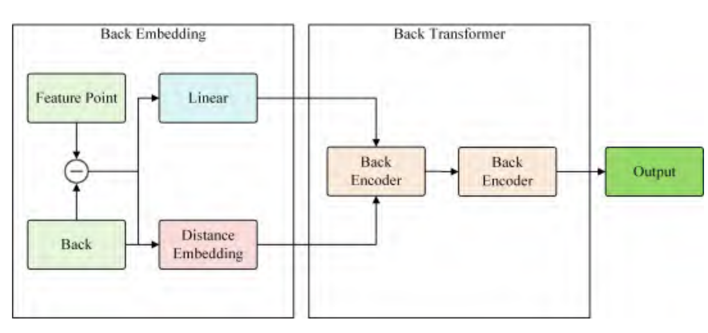
式中：$X_{d1,i}$是每个点到关键点的距离；m为属于M的点。

X_d的第二维度表示为公式（9）。

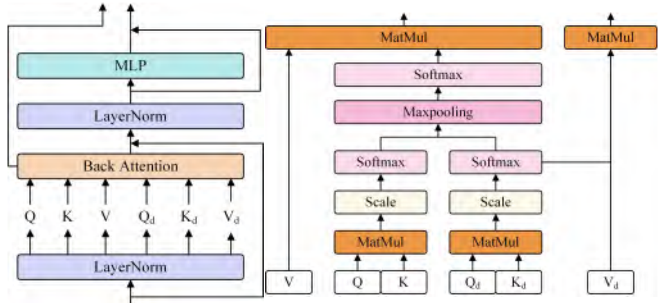
$$X_{d_{2,i}} = m_i - g_{i2} = \sqrt{\sum_{j=1}^{3}\left(M_{j,i} - G_{j,i}\right)^2} \tag{9}$$

式中：$X_{d_{2,i}}$是每个点到质心的距离；m为属于M的点；g为质点。

加入质心和全局特征点偏移的X_d为神经网络中引入了关于猪体大小和姿态的信息。当猪的大小不同时，两种偏移值的整体都会相应的产生变化。同时，X_d也蕴含了猪的运动姿态信息。



a. Back Embedding & Back Transformer



b. Back Encoder　　c. Back-Attention

图10 Back Embedding & Back Transformer结构

Fig. 10 Structures of Back Embedding & Back Transformer

猪的运动平面在运动解剖学中将分为冠状面、矢状面和横断面，其中冠状面平行于地面，高度位于猪脊柱矢状面垂直于地面、平行于猪脊柱；横断面为同时垂直于地面与猪脊柱的平面，采集过程中发生的主要运动为位于冠状面与矢状面的运动[18]。

当姿态发生冠状面的变化时，如猪左右扭头和左右转身，两侧的质心偏移距离的变化会比较剧烈，若将此时猪背假设为弧形，质心大幅向弧度的中心移动，由于质心的初始位置为接近猪背部中心点，所以猪前侧和后侧与质心的距离减小，中心处与质心的距离增加，同时两侧的减小量比中心的增加量要更大；而全局特征点在学习过程中会发生沿猪的背脊走向移动的整体倾向，此时偏移距离变化较小的部分为移动较少的部分，隐式地表述猪发生运动的关节位置。

当猪姿态发生矢状面变化时，如猪低头、抬头、弓背，由于网络中只使用了猪背部点云，因此对于中心的质心偏移量影响会更大，此时的前后侧偏移距离也是具有前侧和后侧与质心距离减小，中部点与质心距离增加的趋势，同时两侧的减小量比中心的增加量更小；全局特征点在此时会向中心收缩，导致全局偏移量和质心偏移量的相似度提高。

最终，X_d可以在其张量空间上通过距离隐式含有猪的不同姿势和大小信息。

Back Transformer通过两层连续的编码器组成。编码器结构改进为同时输出提取到点云的特征和距

离特征，同时注意力部分使用了改进的back attention结构。back attention的设计如公式（10）和公式（11）所示。

$$Attn_p\left(X_p, X_d\right) = sm\left(mp\left(sm\left(\frac{\left(X_pW_{pq}\right)\left(X_pW_{pk}^T\right)}{\sqrt{D_p}}\right), sm\left(\frac{\left(X_dW_{dq}\right)\left(X_dW_{dk}^T\right)}{\sqrt{D_{pd}}}\right)\right)\right) \times \left(X_pW_v\right) \tag{10}$$

$$Attention_v\left(X_d\right) = softmax\left(\frac{\left(X_dW_{dq}\right)\left(X_dW_{dk}^T\right)}{\sqrt{D_{pd}}}\right) \times \left(X_pW_v\right) \tag{11}$$

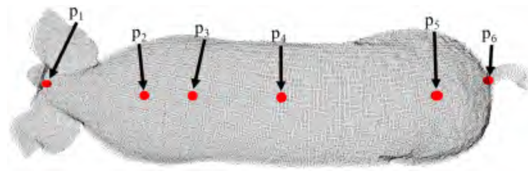
式中：X_p，X_d分别为使用线性层和双重距离嵌入得到的特征；W为权值矩阵；D为维度。

相比self attention结构，back attention加入了距离信息作为额外的注意力来源。与edge attention中的叠加自注意权值的做法不同的是，back attention中对X_p和X_d的自注意力权值进行了按位max pooling的操作，采用max pooling的原因是X_d蕴含了足够的信息，使其产生的自注意力具备较强代表性，所以通过max pooling后得到的自注意力权值可以同时兼顾对求解关键点需要的全局信息、相对位置信息和姿态信息。

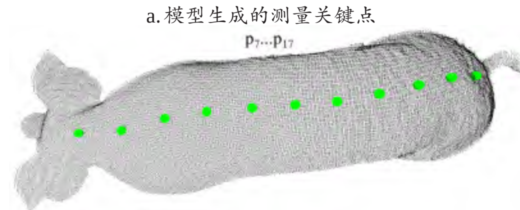
在Back Transformer提取完成特征后，通过两层的MLP得到17个偏移量，再将偏移量与全局特征点求和，模型最终生成17个点（图11），其中6个测量关键点p_1～p_6和11个背脊走向点p_7～p_{17}。

3） 标记与损失函数

本研究中每个猪背部点云标记了6个测量关键点和多个背脊走向点，如图12所示，测量关键点从头至尾依次为：体长测量起点，前体高、肩宽关键点，胸围关键点，腹围、腹宽关键点，后体高、臀宽关键点，以及体长测量终点。



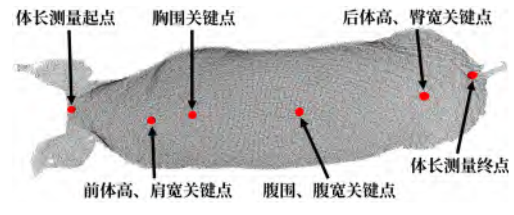
a. 模型生成的测量关键点



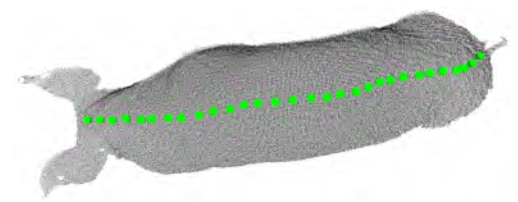
b. 模型生成的背脊走向点

图 11　Pig Back Transformer模型生成的关键点

Fig. 11　Key points generated by Pig Back Transformer



a. 人工标记的测量关键点



b. 人工标记的背脊走向点

图 12　人工标记的测量关键点和背脊走向点

Fig. 12　Human marked measurement key points and spine direction points

损失函数包含三个目标。第一个目标是最小化生成的点和人工标注的测量关键点的距离；第二个目标是按顺序生成尽可能近的脊背点；第三个目标是降低相邻生成点间距的方差，缩小方差可以使生成点的分布更加均匀，均匀的分布可以使生成点对背脊线的拟合程度更高，提高体尺计算精度。

损失函数如公式（12）～公式（18）所示。

$$Loss = Loss_1 + Loss_2 + Loss_3 \tag{12}$$

$$Loss_1 = \sum_{i=1}^{6} \|k_i - g_i\|_2 \tag{13}$$

式中：k_i表示第i个生成点；g_i表示第i个人工标记的测量关键点；$Loss_1$为生成的6个测量关键点和人工标记的6个测量关键点对应的距离。

$$Loss_2 \sum_{i=7}^{17} (\|k_i - e_i\|_2 + \left(k_i - e_i\right)\left(k_i - f_i\right)\left(1 - \theta_i\right)^4) \tag{14}$$

$$\theta_i = \frac{v_i \cdot u_i}{\|v_i\|_2 \times \|u_i\|_2}, v_i = k_i - k_{i-1}, u_i = k_{i+1} - k_i \tag{15}$$

式中：e_i为k_i与人工标记背脊走向点集合中的最近点；f_i为k_i与人工标记背脊走向点集合中第二近点；$\left\|k_i - e_i\right\|$表示每个点与躯干上最近点的距离，用来约束生成点的精度；$\left(k_i - e_i\right) \times \left(k_i - f_i\right)$表示的是生成的点和躯干的偏差，用来约束生成的点落在Ground Truth走向上的程度；$\left(1 - \theta_i\right)^4$表示的是生成点间连线的方向是否一致，用来约束生成点的顺序。

$$Loss_3 = \left(\bar{k} - d_i\right)^2 \tag{16}$$

$$d_i = \|k_i - k_{i+1}\|_2 \tag{17}$$

$$\bar{k} = \frac{d_i}{T-1} \tag{18}$$

式中：d_i为点k_i和点k_{i+1}间的距离；$\bar{k}$为相邻点之间的平均距离；$\left(\bar{k} - d_i\right)^2$为相邻点间距离的方差，用来约束生成点的均匀性，当点越均匀时相邻点之间的平均距离的方差越小，此时脊背线走向点的分布更加均匀。

1.2.2 猪体尺算法

本研究依据Pig Back Transformer产生的体尺关键点和背脊走向点，设计了对应的求体长、前体高、后体高、肩宽、腹宽、臀宽、胸围、臀围的算法。

开始体尺计算前，首先要将地面调整到xOz平面，计算模型生成点在S_p的最近点集合I，并使用RANSAC（Random sample consensus）算法[19]拟合得到地面方程，将S_p和I根据地面方程旋转，使地面法向量平行于y轴且与xOz平面距离为0。

1）体长关键点如图13所示，红色为头部和尾部关键点，分别位于两耳中心和尾部末端，绿色为背部曲线拟合点。

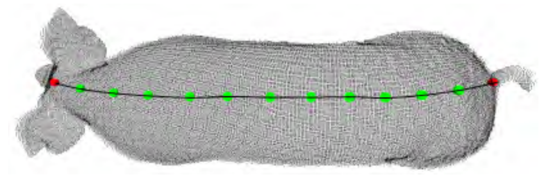


图13 Pig Back Transformer模型的猪体长关键点与拟合曲线

Fig. 13 Key points for pig body length calculating & the fitting curve in Pig Back Transformer model

将头尾的测量关键点p_1，p_6，11个背脊走向点p_7～p_{17}和背部点云都投影到xOy平面上，对头尾关键点和11个背脊走向点的投影进行最小二乘拟合背脊走向曲线，再以2 cm为间隔在曲线上取得曲线和点云投影最近点点集J，将J映射回背部点云得到K，体长为K的点间距之和。

$$Length = \sum_{i=1}^{T-1} \|k_i - k_{i+1}\|_2 \tag{19}$$

2）体高关键点如图14a所示，左侧为前体高测量点，位于肩胛中心，右侧为后体高测量点，位于髋部顶部。



a. 俯视角体高测量点和切片区域

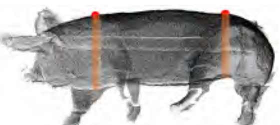
b. 俯视角体宽测量点和切片区域



c. 俯视角胸围测量点和切片区域



d. 俯视角腹围测量点和切片区域





e. 侧视角前后体高测量点

f. 侧视角体宽测量点



g. 侧视角胸围测量点和切片区域



h. 侧视角腹围测量点和切片区域

图14 体尺算法中的猪体高、体宽、胸围、腹围关键点和切片区域

Fig. 14 Key points & slicing area for height, width, chest circumstance and abdomen circumstance in body size calculation algorithm

前体高为前体高关键点处沿垂直于走向曲线做切片得到集合X_{Front}中前10%高点的平均高度：

$$H_{Front} = \max\left(y\left(\boldsymbol{x}_i\right)\right), \boldsymbol{x} \in \boldsymbol{X}_{Front} \tag{20}$$

后体高为后体高关键点处沿垂直于走向曲线做切片得到集合X_{Back}中前10%高点的平均高度：

$$H_{Back} = \max\left(y\left(\boldsymbol{x}_i\right)\right), \boldsymbol{x} \in \boldsymbol{X}_{Back} \tag{21}$$

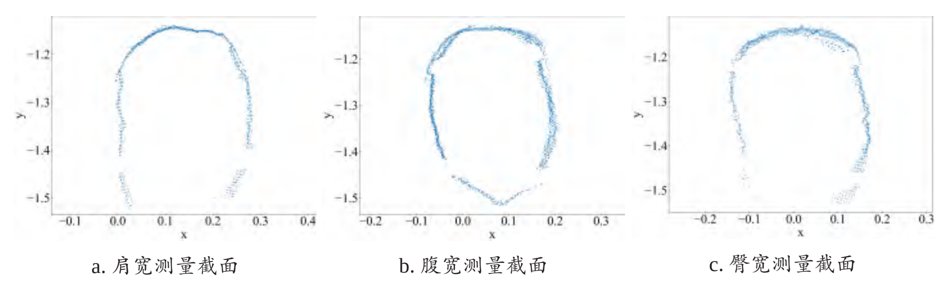
3）体宽关键点如图14b所示，从左到右依次为肩宽、腹宽、臀宽测量点。

肩宽计算方法为在肩宽关键点处沿垂直于走向曲线做切片得到集合C，并旋转切片集合，使切片集合平行于xOy平面，再等间距地分割切片得到n个集合$C_1 \cdots C_n$，在每个集合中存在$W_i = \max\left(x\left(C_i\right)\right) - \min\left(x\left(C_i\right)\right)$，前体宽即为$\max\left(W_i\right)$。测量截面如图15a所示。

腹宽计算方法为在腹宽关键点处使用与肩宽相同算法得到。测量截面如图15b所示。

臀宽计算方法为在臀宽关键点处使用与肩宽相同算法得到。测量截面如图15c所示。

4）胸围算法为在图14c所示的胸围关键点处沿垂直于走向曲线做切片得到点集合$\boldsymbol{D}$，用体素降采样降低点云密度，再旋转切片集合使其平行于xOy平面，将$\boldsymbol{D}$投影到xOy平面，用统计方法去除切片中的腿部部分，再将投影后的点转换为极坐标，通过拟合极坐标曲线进行近似补全，最后将极坐标转换回平面直角坐标系得到如图16所示的闭



a. 肩宽测量截面　　b. 腹宽测量截面　　c. 臀宽测量截面

图 15　猪体宽计算过程中的肩宽、腹宽、臀宽的测量截面图

Fig. 15　Slice for shoulder, abdomen and hip width calculation in width calculation process

合曲线 l_{chest}，使用微元法计算闭合曲线 l_{chest} 的长度 C_{chest} 即胸围计算结果。

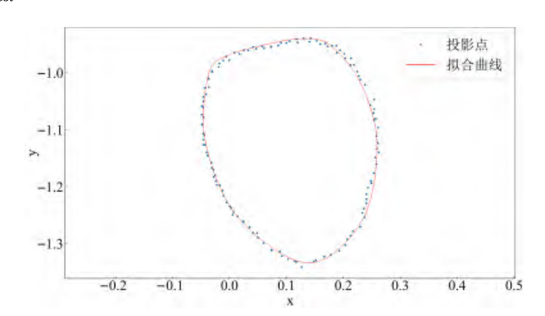


图 16　胸围计算过程中的胸围曲线拟合结果

Fig. 16　Fitting results of chest circumference

5）腹围算法为在图14d所示的腹围关键点处做与胸围相同算法，得到如图17所示的闭合曲线 l_{abd}，使用微元法计算闭合曲线 l_{abd} 的长度 C_{abd} 即为腹围计算结果。

2　实验结果与分析

实验使用的硬件为 AMD Ryzen Threadripper PRO 3995WX CPU 和 RTX 3090 GPU，操作系统为 Linux，神经网络框架为 Pytorch。

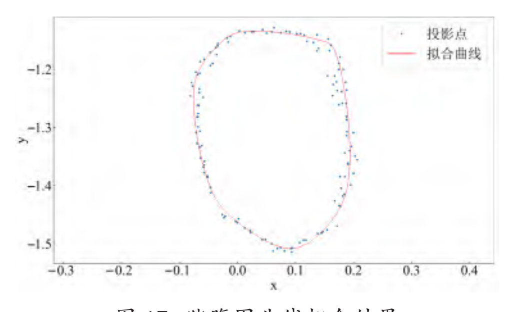


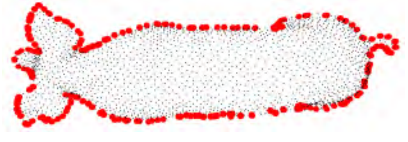
图 17　猪腹围曲线拟合结果

Fig. 17　Fitting curve of pig's abdomen circumference

2.1　边缘提取

边缘提取实验前先对点云进行体素大小为0.01的体素降采样，再使用最远点降采样将点云降采样至2 048个点提升计算速度。

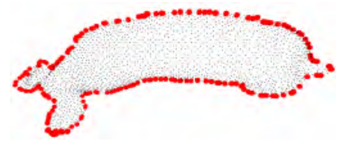
边缘提取效果如图18所示，可以看出本研究提出的边缘提取算法提取出了猪背部点云的边缘，泛化性较强，可以用于不同姿态下的猪背部点云。



a. 标准姿态的边缘提取结果



b. 部分缺失的边缘提取结果



c. 大幅弯曲姿态的边缘提取结果

图 18　顶部视角猪背部点云的边缘提取结果

Fig. 18　Results of edge detection for pig back's point cloud captured from top angle

但是，从图19所示的侧视角可以看到，点云边缘的曲率变化剧烈、连续性较差，在使用Shared MLP对边缘点云提取特征时，如果不对点云边缘进行预处理，会导致训练过程中边缘点中提取的特征不具有代表性，影响网络的性能和训练时的稳定性，此问题出现的原因是Shared MLP会对输入的点云中的所有点进行相同的线性映射，所以当样本相似度较低时有可能出现欠拟合现象。

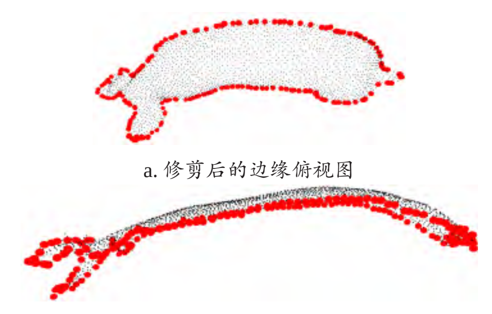
而在经过修剪后的边缘提取结果如图20所示，连续性更强，边缘更平整，受猪的姿态影响更小，在神经网络训练时性能更好，稳定性更强。



图 19　顶部视角猪背部点云的边缘提取结果的侧面

Fig. 19　Side view of the result of edge detection for pig back's point cloud captured from top angle

在后续的神经网络训练实验中进一步量化修剪



a. 修剪后的边缘俯视图

b. 修剪后的边缘侧视图

图 20 顶部视角猪背部点云的修剪后的边缘提取效果

Fig. 20 Edge detection result after trimmed for pig back's point cloud captured from top angle

与否的稳定性和性能差异。

2.2 关键点生成

在关键点生成实验中，Pig Back Transformer模型参数量为10 M，数据集为前文介绍的人工标记的810组数据，通过旋转、轴对称操作将数据增强到4 050个点云。

2.2.1 与其他模型的对比

实验所用模型性能如表1所示。表1中数据显示出实验所用的模型中，PointNet++的生成点均匀度较好，但测量关键点和背脊走向点的平均偏差较高；Point Transformer V2的测量关键点和背脊走向点生成精度高于PointNet++，均匀度和PointNet++持平；Point Cloud Transformer的精度和均匀度都略差；OctFormer的关键点和背脊走向点精度较高，均匀度表现为实验中最好；PointTr的精度和均匀度都较高；Pig Back Transformer精度表现最好，均匀度较高。

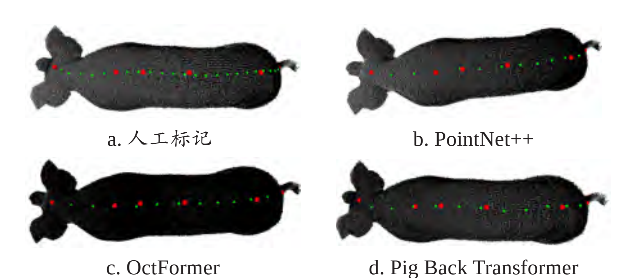
关键点与背脊走向点生成效果如图21所示，PointNet++的生成点均匀度较好，但关键点和走向点的平均偏差较高，OctFormer的关键点和走向点平均偏差都较低，均匀度表现为实验中最好，Pig Back Transformer精度表现最好，均匀度满足测量需求。

表 1 不同模型的关键点拟合结果

Table 1 Key points generated by different models

模型	Loss	关键点平均偏差	点间距方差	参数量/M
PointNet++	0.061	0.009	0.001 5	**8.7**
Point Transformer V2[20]	0.052	0.008	0.001 5	11.0
Point Cloud Transformer[21]	0.074	0.012	0.002 1	10.5
OctFormer[22]	0.037	0.005	**0.001 0**	11.8
PoinTr[23]	0.082	0.013	0.001 8	10.1
Pig Back Transformer(无边缘提取模块)	0.091	0.015	0.001 5	9.8
Pig Back Transformer	**0.013**	**0.002**	0.001 1	10.3

注：加粗为性能最好。



a. 人工标记 b. PointNet++

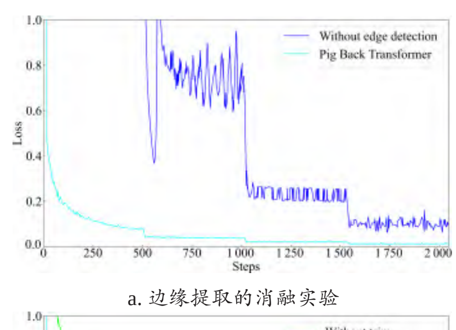
c. OctFormer d. Pig Back Transformer

图 21 Point++，OctFormer，Pig Back Transformer关键点和猪背脊走向点拟合效果和人工标记对比图

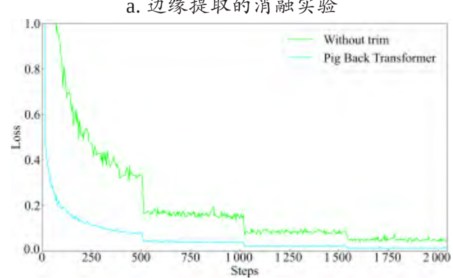
Fig. 21 Key points & pig spine direction points generated by PointNet++，OctFormer，Pig Back Transformer compare with manual labeling

2.2.2 消融实验

消融实验如图22所示。在去除掉边缘提取模块，并将自注意力机制修改为传统的Transformer后，直接使用背部点云进行关键点生成时，稳定性和拟合效果都大幅下降。在去除掉边缘修整处理后，点云的训练速度和拟合效果都发生了下降，并且在训练时的稳定性也受到了很大的影响。



a. 边缘提取的消融实验



b. 点云修剪的消融实验

图 22 边缘提取和点云修剪的消融实验

Fig. 22 Ablation experiment of edge detection and point cloud trim

生成关键点精度的任务上，Pig Back Transformer相对其他点云处理网络得到了提高，并且达到了体尺计算所需精度。

2.3 体尺计算

2.3.1 与人工测量结果对比

选取了第一次采集的20头猪体，使用每头猪10个点云计算体尺数据并取平均值作为体尺计算结果并与人工结果进行对比（表2）。

表 2 Pig Back Transformer模型体尺计算结果与人工体尺测量的误差

Table 2 Error comparing body size calculated by Pig Back Transformerwith human measured result

体尺类型	平均相对误差/%	平均误差/cm	最大误差/cm	误差标准差
体长	0.63	0.729	1.982	0.079
前体高	1.85	1.132	2.264	2.199
后体高	2.38	1.610	3.395	0.678
肩宽	2.68	0.832	2.264	2.348
腹宽	1.22	0.369	0.851	0.435
臀宽	5.46	1.747	3.608	1.194
胸围	4.02	4.200	3.170	2.910
腹围	3.95	4.585	2.966	2.831

结果表明，体长的相对和平均误差结果较好，但由于猪体变化导致的最大误差仍然较高。当猪体发生冠状面变化时，体表皮肤会发生拉伸，对此时采集到的点云进行体长计算会产生较大的误差。

前体高平均误差和最大误差都好于后体高，但标准差比后体高更高，产生这种现象的原因是算法设计过程中未考虑猪在行走过程中后体高测量处出现肌肉隆起，导致整体存在较高的误差。前体高平均误差较低，但是偶有发生在行走过程中低头的现象，导致误差不稳定，故前体高有较高的误差标准差。

腹宽的平均相对误差、平均误差、最大误差和误差标准差结果相对肩宽和臀宽都更好，产生这种现象的原因有两点：一是猪体发生姿态变化时腹宽相对肩宽与臀宽发生的变化会更小；二是肩宽与臀宽的截面截取算法在应对姿态改变时需要优化，由于行走时腿部连线并不垂直于背脊线，所以算法所截平面与实际物理平面存在误差。

胸围与腹围的测量结果都较差，是由于两侧相机视角与地面的相对高度较高，部分点云出现了的较大范围的腹部和胸部空洞，这些空洞导致通过极坐标最小二乘拟合算法得到的胸部、腹部曲线与猪体实际的物理形状相差较大，因此产生了较大的误差。

2.3.2 猪体姿态剧烈变化对体尺的影响

由于猪体姿态会影响点云的体尺测量结果，本研究从第一次采集的20头猪体中选取了背脊走向曲率最高的10头猪，每头猪10个点云进行同一头猪不同姿势的体尺分布实验，以测试本模型在不同姿态下体尺测量的稳定性。计算结果如表3和表4所示，分布结果如图24所示，在进行体长计算时，本模型结果较为集中，浮动范围在3 cm以内，在极端情况下因运动幅度过大会产生少数的异常值，在进行腹宽计算时，由于腹宽相对体长基数较小，因此本模型结果在数值上会更为集中，浮动范围在1 cm左右，相对浮动则都在3%左右。

表 3 猪姿态剧烈变化时的体长测量结果

Table 3 Body length results when pig's posture changed significantly

编号	1	2	3	4	5	6	7	8	9	10
人工测量结果/cm	102.0	100.0	111.0	120.0	108.0	115.0	115.0	116.0	127.0	129.0
结果1/cm	102.5	99.1	111.1	121.1	108.4	114.1	115.0	116.3	127.4	128.9
结果2/cm	102.2	100.1	112.1	119.5	108.1	114.4	114.0	116.0	128.9	130.4
结果3/cm	101.1	98.8	109.8	119.0	108.9	113.3	113.8	116.5	128.8	128.6
结果4/cm	102.1	109.0	109.4	118.4	107.2	114.7	114.6	116.2	127.6	129.4
结果5/cm	101.5	101.1	109.9	118.3	109.0	115.1	114.2	116.0	128.4	128.6
结果6/cm	100.1	101.2	109.8	119.2	108.0	115.5	115.8	116.4	128.6	130.7
结果7/cm	102.0	100.8	109.8	119.3	108.5	115.0	115.6	114.3	128.7	130.8
结果8/cm	103.9	100.9	109.0	119.3	109.5	115.5	115.2	115.4	128.8	129.2
结果9/cm	102.9	100.5	109.9	119.0	108.2	115.2	115.6	117.4	128.2	130.0
结果10/cm	102.8	99.4	111.0	119.4	108.8	114.9	115.9	116.9	128.6	129.4

3 讨　论

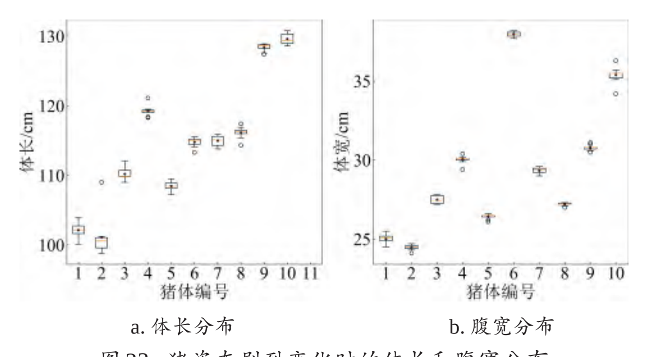
3.1 Pig Back Transformer实验中的不足

Pig Back Transformer的实验中存在两点不足。第一是未使用公共数据集进行关键点生成实验。公开数据集的多样性能对模型的泛化性进行更充分的测试，同时使用公开数据集进行实验结果会更具有公信力，但两个问题导致了无法使用公开数据集进行实验：首先，目前公开数据集中的关键点均为关

表 4 猪姿态剧烈变化时的腹宽测量结果

Table 4 Abdomen width results when posture changed significantly of pig

编号	1	2	3	4	5	6	7	8	9	10
人工测量结果/cm	25.0	25.0	27.0	30.0	27.0	38.0	29.0	27.0	30.0	35.0
结果1/cm	25.1	24.5	27.8	30.1	26.5	38.1	29.5	27.2	31.0	35.5
结果2/cm	25.1	24.4	27.5	29.4	26.4	38.0	29.4	27.3	30.8	35.1
结果3/cm	25.2	24.4	27.4	30.2	26.6	38.0	29.4	27.1	30.8	35.1
结果4/cm	25.3	24.3	27.6	30.1	26.5	38.0	29.4	27.2	30.7	35.5
结果5/cm	25.5	24.5	27.8	30.1	26.2	37.8	29.5	27.0	31.1	36.3
结果6/cm	24.9	24.1	27.8	29.9	26.5	37.8	29.0	27.2	30.8	35.5
结果7/cm	24.9	24.5	27.4	30.0	26.1	37.7	29.0	27.2	30.5	35.7
结果8/cm	24.5	24.6	27.2	30.4	26.5	38.1	29.3	27.2	30.6	34.2
结果9/cm	25.0	24.6	27.2	30.1	26.5	38.2	29.6	27.3	30.8	35.5
结果10/cm	24.5	24.7	27.2	30.0	26.4	38.1	29.2	27.3	30.7	35.5



a. 体长分布　　b. 腹宽分布

图23 猪姿态剧烈变化时的体长和腹宽分布

Fig. 23 The distributed of body length and abdomen width when pig's posture changed significantly

键点配对任务，与本研究所需的提取测量关键点任务不同，无法作为本研究的性能指标，也无法进行迁移；其次，本研究采用了点云边缘提取，因为观察到猪背部点云的边缘蕴含的几何信息对整个猪背部点云具有代表性，而公开数据集则不符合此情况。第二是由于边缘提取部分与编码器设计耦合程度太高，导致边缘提取模块和注意力机制的消融实验无法独立进行。

3.2 Pig Back Transformer模型的不足和改进方向

本模型的不足之处有两点。第一是自动化程度仍存在提高空间，由于构建训练集时采用人工剔除掉不完整点云，若直接将本模型应用在生产环境中，便需要人工确定猪体点云是否超出摄像机采集范围。第二是对点云形状的兼容范围不够广，由于采用了边缘提取作为部分输入，算法的适用范围为某一视角的点云存在足够提取出稳定的全局关键点的几何信息。

模型存在四个可改进方向。第一，通过深度图和同时拍摄的二维图像进行关键点检测。深度图在像素上与点云一一对应，并且自身存在结构化信息，可以使用二维计算机视觉方法如YOLO等进行关键点检测，定位到像素点后再映射到点云中的对应点即可进行体尺计算，但是点云作为经过相机内参转换的深度图蕴含了相机内参，而使用深度图和二维图像进行目标检测则缺少这部分信息。第二，利用深度图与点云一一对应的关系，将降噪和去除背景后的点云结构化并利用机构化信息提高特征提取性能。第三，可以通过投影的方法提升模型的适用范围，将完整点云分别投影到不同视角，如投影到xOy，yOz，xOz平面，再进行不同视角下的边缘提取，最后通过多视角下的边缘点集输入模型生成关键点。第四，设计算法检测猪体点云是否超出摄像机采集范围，自动构建模型训练所需数据集，同时提高生产环境中的自动化程度。

4 结　论

本研究设计的关键点生成模型Pig Back Transformer有较高的关键点生成精度，相对其他的点云处理网络有较大提升，关键点平均偏差为0.2 cm，解决了猪体尺测量中的关键点定位和背脊走向拟合问题。依托体尺测量关键点构建的体尺算法在体长、体高、体宽测量上获得了较高精度，体长的平均相对误差为0.63%，相对其他研究方法有较大提升。同时，提出的算法能较好地处理动物运动过程中体态的变化，在姿态剧烈变化下体尺的分布较为集中，只有少量异常值产生，具有较好的泛化性，可以改善采集过程中因姿态变化产生的体尺测量误差。

综上所述，本研究设计的种猪体点云关键点生成模型，提高了关键点生成精度。根据关键点和背部走向设计的体尺计算算法，降低了无接触测量的误差。得到的高精度无接触体尺测量结果可以为猪

育种工作提供坚实的数据基础。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 江杰, 周丽娜, 李刚 . 基于机器视觉的羊体体尺测量[J]. 计算机应用, 2014, 34(3): 846-850, 887.
JIANG J, ZHOU L N, LI G. Sheep body size measurement based on computer vision[J]. Journal of computer applications, 2014, 34(3): 846-850, 887.

[2] ZHANG A L, WU B P, TANA WUYUN C, et al. Algorithm of sheep body dimension measurement and its applications based on image analysis[J]. Computers and electronics in agriculture, 2018, 153: 33-45.

[3] YANG G Y, XU X S, SONG L, et al. Automated measurement of dairy cows body size via 3D point cloud data analysis[J]. Computers and electronics in agriculture, 2022, 200: ID 107218.

[4] LINDEBERG T. Scale invariant feature transform[J]. Scholarpedia, 2012, 7(5): ID 10491.

[5] SALAU J, HAAS J H, JUNGE W, et al. Feasibility of automated body trait determination using the SR4K time-of-flight camera in cow barns[J]. Springerplus, 2014, 3: ID 225.

[6] 叶文帅, 康熙, 贺志将, 等 . 基于深度图像的多姿态肉牛体尺自动测量方法[J]. 智慧农业(中英文), 2022(4): 144-155.
YE W S, KANG X, HE Z J, et al. Automatic measurement of multi-posture beef cattle body size based on depth image[J]. Smart agriculture, 2022(4): 144-155.

[7] RUCHAY A, KOBER V, DOROFEEV K, et al. Accurate body measurement of live cattle using three depth cameras and non-rigid 3-D shape recovery[J]. Computers and electronics in agriculture, 2020, 179: ID 105821.

[8] SHI W, DAI B, SHEN W, et al. Automatic estimation of dairy cow body condition score based on attention-guided 3D point cloud feature extraction[J]. Computers and electronics in agriculture, 2023, 206: ID 107666.

[9] HU H, YU J, YIN L, et al. An improved PointNet++ point cloud segmentation model applied to automatic measurement method of pig body size[J]. Computers and electronics in agriculture, 2023, 205: ID 107560.

[10] YIN L, ZHU J, LIU C, et al. Point cloud-based pig body size measurement featured by standard and non-standard postures[J]. Computers and electronics in agriculture, 2022, 199: ID 107135.

[11] CHARLES R Q, HAO S, MO K C, et al. PointNet: deep learning on point sets for 3D classification and segmentation[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2017.

[12] ZHAO H, JIANG L, JIA J, et al. Point Transformer[EB/OL]. arXiv: 2012.09164, 2021.

[13] PARK C, JEONG Y, CHO M, et al. Fast Point Transformer[EB/OL]. arXiv: 2112.04702, 2022.

[14] YU X M, TANG L L, RAO Y M, et al. Point-BERT: Pre-training 3D point cloud transformers with masked point modeling[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2022.

[15] 尹令, 蔡更元, 田绪红, 等 . 多视角深度相机的猪体三维点云重构及体尺测量[J]. 农业工程学报, 2019, 35(23): 201-208.
YIN L, CAI G Y, TIAN X H, et al. Three dimensional point cloud reconstruction and body size measurement of pigs based on multi-view depth camera[J]. Transactions of the Chinese society of agricultural engineering, 2019, 35 (23): 201-208.

[16] LIN R H, HU H, WEN Z K, et al. Research on denoising and segmentation algorithm application of pigs' point cloud based on DBSCAN and PointNet[C]// 2021 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor). Piscataway, New Jersey, USA: IEEE, 2021: 42-47.

[17] QI C R, YI L, SU H, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space[EB/OL]. arXiv: 1706.02413, 2017.

[18] HOLTZMAN N G, YAKUBOV D J. Fetal Pig Dissection Manual (BIOL 105) [M]. New York: New York Cuny Queens College, 2019.

[19] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[M]// Readings in Computer Vision. Amsterdam: Elsevier, 1987: 726-740.

[20] WU X, LAO Y, JIANG L, et al. Point transformer V2: Grouped vector attention and partition-based pooling[EB/OL]. arXiv: 2210.05666, 2022.

[21] GUO M H, CAI J X, LIU Z N, et al. PCT: Point cloud transformer[J]. Computational visual media, 2021, 7(2): 187-199.

[22] WANG P S. OctFormer: Octree-based transformers for 3D point clouds[J]. ACM transactions on graphics, 2023, 42(4): 1-11.

[23] YU X M, RAO Y M, WANG Z Y, et al. PoinTr: diverse point cloud completion with geometry-aware transformers[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, New Jersey, USA: IEEE, 2021.

Pig Back Transformer: Automatic 3D Pig Body Measurement Model

WANG Yuxiao[2], SHI Yuanyuan[1], CHEN Zhaoda[1], WU Zhenfang[2,3,4], CAI Gengyuan[2,3,4], ZHANG Sumin[1], YIN Ling[1,2,4*]

(1. *College of Mathematics and Informatics*, *South China Agricultural University*, *Guangzhou* 510642, *China*; 2. *National Engineer-*

ing Research Center For Breeding Swine Industry, Guangzhou 510642, *China*; 3. *College of Animal Science, South China Agricultural University, Guangzhou* 510642, *China*; 4. *National Key Laboratory of Pig and Poultry Breeding Industry, Guangzhou* 510640, *China*)

Abstract:
[Objective] Nowadays most no contact body size measurement studies are based on point cloud segmentation method, they use a trained point cloud segmentation neural network to segment point cloud of pigs, then locate measurement points based on them. But point cloud segmentation neural network always need a larger graphics processing unit (GPU) memory, moreover, the result of the measurement key point still has room of improvement. This study aims to design a key point generating neural network to extract measurement key points from pig's point cloud. Reducing the GPU memory usage and improve the result of measurement points at the same time, improve both the efficiency and accuracy of the body size measurement.
[Methods] A neural network model was proposed using improved Transformer attention mechanic called Pig Back Transformer for generating key points and back orientation points which were related to pig body dimensions. In the first part of the network, it was introduced an embedding structure for initial feature extraction and a Transformer encoder structure with edge attention which was a self-attention mechanic improved from Transformer's encoder. The embedding structure using two shared multilayer perceptron (MLP) and a distance embedding algorithm, it takes a set of points from the edge of pig back's point cloud as input and then extract information from the edge points set. In the encoder part, information about the offset distances between edge points and mass point which were their feature that extracted by the embedding structure mentioned before incorporated. Additionally, an extraction algorithm for back edge point was designed for extracting edge points to generate the input of the neural network model. In the second part of the network, it was proposed a Transformer encoder with improved self-attention called back attention. In the design of back attention, it also had an embedding structure before the encoder structure, this embedding structure extracted features from offset values, these offset values were calculated by the points which are none-edge and down sampled by farthest point sampling (FPS) to both the relative centroid point and model generated global key point from the first part that introduced before. Then these offset values were processed with max pooling with attention generated by the extracted features of the points' axis to extract more information that the original Transformer encoder couldn't extract with the same number of parameters. The output part of the model was designed to generate a set of offsets of the key points and points for back direction fitting, than add the set offset to the global key point to get points for pig body measurements. At last, it was introduced the methods for calculating body dimensions which were length, height, shoulder width, abdomen width, hip width, chest circumference and abdomen circumference using key points and back direction fitting points.
[Results and Discussions] In the task of generating key points and points for back direction fitting, the improved Pig Back Transformer performed the best in the accuracy wise in the models tested with the same size of parameters, and the back orientation points generated by the model were evenly distributed which was a good preparation for a better body length calculation. A melting test for edge detection part with two attention mechanic and edge trim method both introduced above had being done, when the edge detection and the attention mechanic got cut off, the result had been highly impact, it made the model couldn't perform as well as before, when the edge trim method of preprocessing part had been cut off, there's a moderate impact on the trained model, but it made the loss of the model more inconsistence while training than before. When comparing the body measurement algorithm with human handy results, the relative error in length was 0.63%, which was an improvement compared to other models. On the other hand, the relative error of shoulder width, abdomen width and hip width had edged other models a little but there was no significant improvement so the performance of these measurement accuracy could be considered negligible, the relative error of chest circumference and abdomen circumference were a little bit behind by the other methods existed, it's because the calculate method of circumferences were not complicated enough to cover the edge case in the dataset which were those point cloud that have big holes in the bottom of abdomen and chest, it impacted the result a lot.
[Conclusions] The improved Pig Back Transformer demonstrates higher accuracy in generating key points and is more resource-efficient, enabling the calculation of more accurate pig body measurements. And provides a new perspective for non-contact livestock body size measurements.

Key words: Pig Back Transformer; 3D point cloud; body size automic measurement; key point positioning; depth camera; self-attention mechanism

Foundation items: General Program of the National Natural Science Foundation of China (32172780), Sub-Project of the National Key R&D Program (2023YFD1300202)

Biography: WANG Yuxiao, E-mail: wangyuxiaojoemob@stu.scau.edu.cn
***Corresponding author:** YIN Ling, E-mail: yin_ling@scau.edu.cn